Computing the history of the family: a question of standards

Schurer, Kevin; Wall, Richard

RICHARD WALL AND KEVIN SCHURER

Computing the History of the Family: A Question of Standards

The present state of research

Over many years the Cambridge Group has built up a valuable collection of censuses and listings and, from these, produced sets of tabular analyses (in manuscript) displaying the family and household structure of various communities. The listing file, as it has come to be called[1], comprises a collection of most known enumerations of English communities prior to 1801 which divided the population into name blocks which can be identified as residential units of some sort although the precise definition of that unit often remains unclear[2]. Holdings of censuses dating from after 1801 and listings of overseas populations are selective, reflecting the individual interests of members of the Cambridge Group. In all, approximately 500 censuses of English communities have been included in the collection[3]; not very many when one considers that pre-industrial England was divided into approximately ten thousand ancient parishes.

Arising from the fact that the collection has taken a number of years to assemble, it is inevitable that the analyses refered to above should vary in accuracy, in the conventions followed (in regard, for example, to the iden-

---

[1] *P. Laslett*: The study of social structure from listings of inhabitants, in: *E.A. Wrigley* (ed.): An Introduction to English Historical Demography, 1966, 160-78. The procedures for analysing these documents have subsequently been substantially revised cf. *P. Laslett*: Mean household size over three centuries, in: *P. Laslett* and *R. Wall* (eds): Household and Family in Past Time, Cambridge, 1972, 127. In addition, much stricter criteria are now employed in the selection of documents, particularly when the research focus is on the composition of the household in terms of the relationships of the members to the household head. See *R. Wall*: The household, demographic and economic change in England, 1650-1970, in: *R. Wall, J. Robin* and *P. Laslett* (eds): Family Forms in Historic Europe, Cambridge, 1983, 496 n9. The unpublished listings analyses are available for consultation in the library of the ESRC Cambridge Group for the History of Population.

[2] These issues are explored in greater detail in the introduction to *Wall, Robin* and *Laslett*, 1983, 6-13.

[3] A full list of the places has appeared in successive issues of the journal *Local Population Studies* beginning in 1968. From Local Population Studies 24 (Spring 1980) the list has been repeated and expanded to include details on the range of information on individuals and households to be found in each document.

tification of households)[4] and in the degrees to which the listing has been fully exploited as fresh tables have been added to the series and old ones redrafted. For some years we have been working on the construction of a set of model tables, similar to those recommended by the United Nations for the presentation of current statistics on the family and household but suitably adapted to reflect the range of persons to be found in households in past societies by comparison with more contemporary societies. The tables, some 85 in all, are divided into various sections according to the particular aspect of the population to be investigated: children, the elderly, servants, relatives, boarders and lodgers, employment and class structure, and migration. Two preliminary sections deal with broader concers such as the distribution of the population by age, sex and marital status and the composition of the household in terms of both kin structure and the balance between producters and consumers. The objectives are threefold: to move discussions of household forms from their obsessive preoccupation with the kin structure of households[5], to recommend standard definitions of critical terms, for example 'family', 'household', 'child', 'relative' etc., and to achieve a minimum level of detail in the relevant tabulations. To the latter end the Cambridge Group, for the cost of computer print-out plus postage[6], will make available

[4] One of the principal difficulties is whether to count lodgers and their families as constituting seperate households. The definition of the household favoured by the Cambridge Group is to see lodgers as constituting appendages to the household (cf. the introduction by *P. Laslett* to: Household and Family in Past Time, 1972, 34-9) but this in turn gives rise to a further problem in that strictly speaking the situation in which a lodging group contains children but the main part of the household does not, constitutes a houshold without children. However, there is reason to think that this decision was not applied consistently as listings were processed by a series of researchers at the Cambridge Group.

[5] Cf. the seemingly never-ending debate on the dominance or otherwise of the nuclear and stem family. The cause of the stem family was championed initially by *Berkner*. See: The stem family and the development cycle of the peasant household, in: American Historical Review, 77/2, 398-418. *Laslett*'s reply came eventually in a contribution to *K.W. Wachter, E.A. Hammel* and *P. Laslett* (eds): Statistical Studies of Historical Social Structure, London: Academic Press, 89-111, but the debate rumbles on particularly in the pages of the Journal of Family History. For a recent contribution see *Harvey Smith*: Family and Class: the household economy of Languedoc wine growers, 1830-1870, in: Journal of Family History, 9/1 (Spring 1984), 65 n.1.

[6] A list of the table titles appears in the Appendix to *P. Laslett*: The family and

to interested researchers any or all of the 85 tables in a form suitable for the insertion of their own results.

The tables were initially tested on three small populations: English, Corsican and Russian, selected because it was felt each would be dominated by quite distinct types of family; such indeed proved to be the case. However, the full potential of the model tables lies in the exploration of the interrelationships of household composition with inheritance custom and practice and particular types of economy and with this objective our research will in the future focus on the better documented communities of mid-nineteenth century Britain and on selected overseas communities. The social structure of mid-nineteenth century Britain is surprisingly under-researched despite censuses of high quality in which individuals are identified by name, age, sex, marital status, relationship to household head, occupation and county and parish of birth. One page from the census enumerator's return for the Dorset parish of Corfe Castle in 1851 is reproduced in Figure 1. In making his return the enumerator failed to keep to the guidelines issued by the census authorities in London (cf below Figure 2 for a transcription and commentary on the errors) but the general format is clear and uniform across the country for a given census year. Censuses survive for virtually every place in the country beginning in 1851 and are currently open for public inspection up to 1881[7]. Despite this, informative studies number only a handful[8] and are al-

---

household as work group and kin group: areas of traditional Europe compared, in *Wall, Robin* and *Laslett*: Family Forms in Historic Europe, 1983, 560-3. Further details are available from *Kevin Schurer* at the Cambridge Group, 27 Trumpington Street, Cambridge CB2 1QA.

[7] Censuses of England and Wales have been held decennially since 1801 with the exception of 1941. Returns of the first four censuses survive in some numbers of in local archives but vary in quality. An almost complete set of returns exists for 1841 but no attempt was made to record marital status and relationship to the head of the household while the ages of adults were rounded down to the nearest five year point (i.e. a man aged 62 with a wife aged 59 would have had their ages recorded as 60 and 55 respectively).

[8] *Dennis Mills* and *Carole Pearce* have produced a usefull bibliography which details the subjects covered in studies based on the unpublished enumeration returns of the mid-19th century censuses. This appeared first as: Census Enumerators' Books (1982), Faculty of Social Sciences, The Open University, but has since been extended and updated as a computerised bibliography in a joint project with the Cambridge Group. Further information is available from the Group's Data Editor, 27 Trumpington Street, Cambridge CB2 1QA.

most exclusively concerned with 1851. Research on the censuses of 1861-1881 is totally uncoordinated and is often of dubious value. As for the mapping of European familial 'regions', designed to establish the principle characterists of the East European, North-West European and the Mediterranean 'family', the approach has been to depend on the analysis of a few arbitrarily selected communities, widely scattered in terms of both historical time and geographical location. These are often then formulated by Peter Laslett in the late 1960s[9].

The existence of a set of model tables, however, will not of itself result in a rapid accretion of knowledge so long as the tables have to be completed by hand since the task can last some months even for a moderately sized parish. An obvious solution is to process the census by computer. In the 1970s, Peter Laslett was recommending data exchange on household by means of the ideographic system[10]: now data are available on magnetic tape but this would be no advance if the data cannot be satisfactorily interpreted. The successful interchange of computerised data requires a standard data collection format, procedures for processing and tabulating data on households and families, and, of course, the set of model tables already referred to so that however different the researchers objectives, the results are strictly comparable. These three requirements can now be met.

Standard Input Format

The proposed standard format for data collection is depicted in Figure 2 using as an example the same page from an enumeration return that was represented in Figure 1. The prime objective is to follow as closely as possible both the content and layout of the original document, avoiding both precoding and abbreviation. In other words, one is producing a transcript in machine-readable form and not a series of codes substituting for a transcript. Not only will this speed the input process but it eliminates the inevitable erros when coding accompanies or precedes input. Naturally, some input errors will occur, but the input of text rather than codes makes checking and rectification of fatal erros that much easier. The present example relates to

[9] Cf. introduction by *Laslett* to: *Laslett* and *Wall*: Household and Family in Past Time, 1972, 74-85 and *Laslett*: Characteristics of the western family considered over time, in *Laslett*: Family Life and Illicit Love in Earlier Generations, Cambridge, 1977, 22, 26.

[10] *Laslett*: Introduction, in: *Laslett* and *Wall*: Household and Family in Past Time, 1972, 41-4.

census but any document containing standard information on individuals or families, such as a manor court roll or parish register, could in principle be handled in a similar way by applying the same logic of input[11]. Additions are made to the text to indicate the structure of the document. These involve the insertion of oblique slashes as field delimiters, the insertion of commas to denote subdivisions within fields as, for example, between prename and surname, and a numeric tag to indicate the nature of the information contained within each line. In the present example (see Figure 2) all lines prefixed by the tag 50 relate to a household and all lines prefixed by tag 60 to an individual. Also a plus sign is added to carry over a line of text beyond the maximum 80 characters. All the information available on that individual is recorded while the household line contains the schedule number and the address. As the household line denotes the start of a new household, it must always be present even if there are no schedule numbers and no addresses as is usually the case with pre-industrial listings.

Other types of line have also been defined but occur less frequently. Lines beginning with 70 indicate the end of a house or houseful as we have come to term it[12]. In the census of 1851 this is indicated by a line drawn across the whole page of the enumeration book indicating that the residents of a new house are being listed[13] Comments made by the researcher can be included in lines commencing with 81-89. These lines may also be used to record occasional or semi-structured information such as page headings[14]. Comments are always confined within double parentheses to set them apart

---

[11] See *K. Schurer*: Historical Databases and the Researcher, in: Computers in Genealogy (forthcoming).

[12] Cf. *Laslett*: Introduction in: *Laslett* and *Wall*: Household and Family in Past Time, 1972, 35-6.

[13] Procedures for identifying households in the 1851 enumeration returns are set out by *Michael Anderson* in Standard Tabulation Procedures for the Census Enumerators' Books 1851-1891, In *E.A. Wrigley* (ed.): Nineteenth Century Society, Cambridge, 1972, 134-45. *Peter Tillott* in the same volume points to some of the pitfalls that await the unwary researcher, see: Sources of inaccuracy in the 1851 and 1861 censuses: 82-133.

[14] In Figure 2 the totals of houses, and of males and females as given by the enumerator at the bottom of the page (cf. Figure 1) appear in line 82. Line 89 is used on five occasions to record both errors in the enumeration and corrections by the enumerator or others. This is required since such cases may cause difficulty in the drawing of boundaries between households or the interpretation of characteristics of a particular individual.

from the text of the document. In addition, a series of statements is included at the commencement of each file, labelling and defining the file in terms of its number, time of creation, character (whether for example it is a census or some other document), detailed description of the community to be investigated and the archive reference number or numbers of the document.

Fields are of variable length but recorded in fixed format, that is the fields of information must follow each other in a strict order. In the present case (Figure 2), the order is as follows: name (subdivided into prenames and surname), relationship to household head, marital condition, male age, female age[15], occupation, and birthplace (subdivided into community, county and country hierarchies)[16]. Censuses containing less or more detail on individuals or households or simply presenting the same information in a different order each require a different format. However, so long as the processor or the data is appraised of the content of each field, this should cause no problem. Furthermore, the choice of a limited and an internationally recommended character set means that texts submitted in this form can be read into different makes of computers[17]. Interchange between computers is also facilitated by not allowing physical line length to exceed 80 characters.

The input process

The various stages of data checking, coding and correction are summarised in Figure 3, based directly on our experience at the Cambridge Group in the processing of census data from Britain, Austria and Turkey[18] using a

[15] An individual's sex was not separately recorded. Instead the age was entered as appropriate into a column for males or for females.

[16] The final column of the enumeration schedule (cf. Figure 1) for the recording of a person who was blind, deaf or dumb is almost invariably blank and is not shown in the transcription.

[17] British Standards Institution, Draft for Development: the structure and representation of data for interchange at the application level (DIAL), Part 1, recommendation for syntax and basic principles, DD75, Part 1 Appendix C (1981).

[18] Machine-readable data for ten Welsh communities in pre-processed form corresponding to the public data format, was supplied by Michael Anderson from his nationwide sample of enumeration districts in 1851. The project is described in *M. Anderson* et al., Preparation and analysis of a machine-readable National Sample from the Enumerators' Books of the 1851 Census of Great Britain. A final report to the Social Science Research Council, HR2066 (1980). The public data format and the advantages of this method of data transmission over the data as collected

computer package called the Statistical Analysis System (abbreviated henceforth to SAS). After punching the census a copy of the raw data is sent to store for security and future reference. In Cambridge the store medium is tape (Figure 3, 1-3). Alternatively, data already on magnetic tape may be acquired directly from other researchers. In these circumstances, it is possible to proceed directly to stages 4 and 5, the creation of a SAS dataset and an inventory of all the items occuring in each field. The inventory is compiled in reference to a coding dictionary of relationships, occupations and marital conditions, abstracted from censuses processed earlier or created from partly or fully coded variables assigned by the data collector. Any items not located in the coding dictionary are output in order that the researcher may assign an appropriate code and update the dictionary. On resubmission all codes are by definition located and can by copied from the dictionary onto the main file and, for safety, copied to tape (Figure 3, step 7-8). The encoding and checking (step 6) is a most important element in the input process in that it provides an opportunity to look for solutions for the incompatabilities between codes. The latter for example would arise if a person had been assigned a code indicating that she was a wife but was reported as being unmarried.

The same programmes also output any entries which seem likely to cause problems of one sort or another in the production of the tables. Some of these issues arise because there may be reason to suspect the accuracy of the date, an extensive age gap between husband and wife, or between mother and child for example. Other difficulties stem from the fact that some of the tables require information on relationship between certain members of the household as well as between each member and the household head, and this is not always self-evident from the given relationship. To take a case in point, whether a grandchild of the head lives with his parents as well as his grandparents may on occasion be inferred from the order in which the individuals are listed in the household while in other cases the situation is uncertain as long as a representative of the intervining generation is present of an age to be the parent and there is no conflict in surname. Others may not agree on the interpretation of such difficult cases and for this reason it is important to preserve as much flexibility as possible. This has been achieved

---

(the data collection format) is described by *John Welford*: The establishment of portable interchange formats for genealogical data – can we hope to reach an acceptable standard?, in: Computers in Genealogy, 1/7 (March 1984), 178-87. *Dr. Manfred Thaller* (Göttingen) and *Dr. Ehmer* (Vienna) supplied the data on a suburb of Vienna in 1880 and *Dr. Cem Behar* and *Dr. Alan Duben* provided the data on 19th century Istanbul.

by preserving the old cases along with the new and these can be activated to produce varying estimates of the number of extented and multiple households while the raw data always remain available for reference. Other checks allow the researcher to identify all cases of married people whose spouse was absent on census night and servants and lodgers who shared a surname with the head of the household and ought to be considered as possible relatives instead of, or in addition to, their status as servants or lodgers.

When all the checking and correcting of the files has been completed, and this usually involves a further run through the correction phase to ensure that all difficulties have been considered and no fresh errors introduced (steps 6 and 9), the revised SAS files are copied to store (Figure 3, step 10) to constitute the working data library. This library will be accessed by a further series of programmes to generate the Model Tables.

The production of the tables

The SAS working data library is a partitioned file and consists in fact of fourmembers. SPBASE (cf. Figure 4) contains all the information on individuals while HOUSE contains all the information on the residential location for the houseful and household. Both members are retained primarily for reference purpose. The member SP is developed from SPBASE and contains the variables and their coded values required for the production of the tables whenever the reference point is the individual rather than the household. The member HH fulfills a similar purpose for households. It contains similar information to the SP file but is organised at the household level rather than at that of the individual, i.e. it holds all information on people in the household, and can be used to count the number of children, relatives and conjugal family units per household. These four data library members form the core of a computerised census analysis system (CAMTAB) which has been developed at the Cambridge Group.

An additional element to the CAMTAB system is a procedure which assigns the Laslett-Hammel classification of household structure to each household. This six-fold classification into *solitaries, no family, simple, extended, multiple,* and *indeterminate households*, first outlined in 1972, has been elaborated but remains essentially unchanged in the present formulation[19]. The

[19] See below Figure 5 and cf. *Laslett*: Introduction, In: *Laslett* and *Wall*: Household and Family in Past Time, 1972, 31. The principal differences concern the more detailed specification in the revised table of the no-family and extended family households.

basic unit is the simple family household of no more than one married couple with or without children, or one lone parent-child group. Extended households are formed by the presence in addition of relatives drawn from outside the immediate nuclear family and the household is considered multiple if these relatives in turn form a secondary marital or parent-child group. Less complex households are represented by persons living alone (solitaries) and by groups of related and unrelated people co-residing, none of them married or being a member of a parent-child group. Each household type is then subdivided to show for example whether the solitaries are widowed or unmarried, whether the simple family households consist of married couples with or without offspring or widows or widowers with offspring and whether extended family households are extended upwards, downwards or laterally from the core marital or parent-child group. All these possibilities are spelt out in Table B3 of the model series (see below Figure 5).

The programme classifies households by type first eliminating lodgers, servants and visitors from consideration and then proceding to identify in turn households of solitaries, no family households and simple family households using such information as the number of persons present, their marital status and their relationship to each other and to the head of the household. Much use has been made by researchers of the Laslett-Hammel classification scheme, and the trained eye can almost instantaneously distinguish an extended from a no family or simple household. However, specifying the logic of these decisions has proved considerably more arduous because all the possibilities must be considered. Figure 6, therefore, is limited to an illustration of the procedures to be followed in identifying the first three household types: solitaries, no family households and simple family households. Together with the further sections which cover extended and multiple family households, these procedures constituted the core of the programme that successfully classified all the households in 10 Welsh communities enumerated in 1851 and nineteenth century Vienna and Instanbul (cf note 18). No doubt, it is possible, indeed likely, that some further modifications of the programme may be required if communities experiencing a yet more complex household type were to be analysed.

Conclusion

Considerable attention has been devoted above to Table B3 because the classification of households in terms of their kin structure has in the past been seen as the arbitor of whether different populations share the same family system. However, further research soon established that there were many

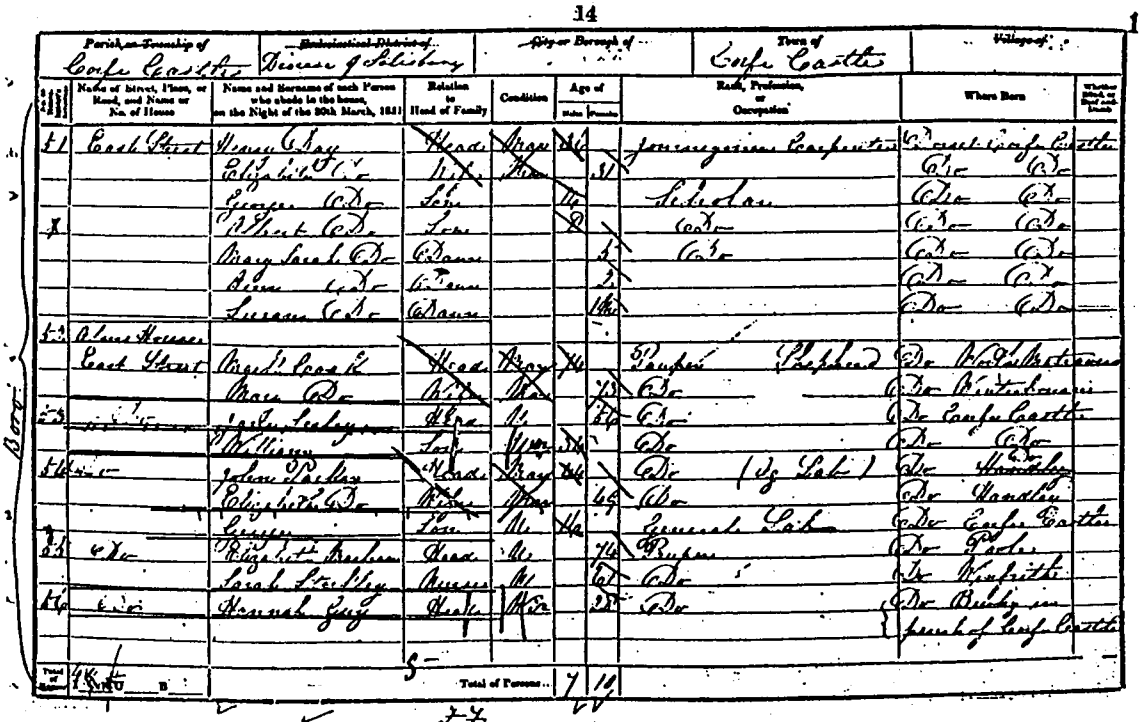
other aspects of the household, for example the presence of servants, the size of the age gap between husband and wife and the changes in the number and type of persons present according to the age of the household head, that were equally deserving of attention[20]. It was to meet such points that the series of model tables was developed and they represent our view of what ought to be involved in a full interpretation of the structure of the households. Nevertheless, it is not to be expected that it will prove possible or even desirable to complete all the tables in every community that is investigated. Nor is it possible to circumvent all the problems of analysing families and households at a point in time when it is well known that their structure is subject to continual change. The process of following individuals, families or households from census to census requires quite a different methodology, and one that is not yet fully developed[21]. Even in terms of the snapshot picture of the household, the set of eighty-five tables may well be insufficiently comprehensive.

Only two claims in their support have been advanced here. The first is that they provide a wide range of standard information on the family and household which individual scholars can consult to place any particular community study in appropriate context. Secondly, the tables serve notice of our intention to end the excessive attention devoted to the kin structure of the household and the tendency to view it in isolation from other features of the society such as employment patterns. In the future, therefore, we will be looking not just at the kin structure of household but whether the more complex households were more or less likely than other types of household to contain servants or inmates. The role of the household as an economic unit will be stressed with a variety of analyses of the size, age and sex composition of its resident labour force. Howeever, it is far from our wish that such lines of enquiry should totally replace the investigation of the kin structure of households. On the contrary, it is the precise relationship between this kin structure and the nature of the local labour market that will be the focus of attention.

---

[20] See especially *Wall*: The composition of households in a population of six men to ten women: south-east Bruges in 1814, In: *Wall, Robin* and *Laslett*: Family Forms in Historic Europe, 1983, 460-72.

[21] Cf. *Wall*: Introduction, In: *Wall, Robin* and *Laslett*: Family Forms in Historic Europe, 1983, 4, commenting on the different approaches of *Reinhard Sieder* and *Michael Mitterauer* and *Luc Danhieu* in their respective contributions to the same volume.



14

140

Parish or Township of: Corfe Castle &c. | Ecclesiastical Division of: Diocese of Salisbury | City or Borough of: | Town of: Corfe Castle | Village of:

Houses uninhabited	Houses building	Name of Street, Place, or Road, and Name or No. of House	Name and Surname of each Person who abode in the house, on the Night of the 30th March, 1851	Relation to Head of Family	Condition	Age of Males	Age of Females	Rank, Profession, or Occupation	Where Born	Whether Blind or Deaf and Dumb
41		Castle Street	Henry Clay	Head	Mar	34		Journeyman Carpenter	Dorset Corfe Castle	
			Elizabeth Do	Wife	Mar		31		Do Do	
			George Do	Son		14		Scholar	Do Do	
X			Albert Do	Son		9		Do	Do Do	
			Mary Sarah Do	Daur			5	Do	Do Do	
			Ann Do	Daur			3		Do Do	
			Susan Do	Daur			1 Mo		Do Do	
42		Alms Houses								
		Castle Street	Mary Ann Cook	Head	Wid		76	Pauper (Parlour)	Do Wool Bastonford	
			Mary Do	Dau	Mar		43	Do	Do Winterborne	
43			Jane Jo Ledger	Head	M		62	Do	Do Corfe Castle	
			William	Son	Wid		24	Do	Do Do	
44			John Parker	Head	Mar	36		Do (Jg Lab)	Do Wambley	
			Elizabeth Do	Wife	Mar		44	Do	Do Wambley	
			George	Son		14		Errand Lab	Do Corfe Castle	
45		Do	Edgar Ann Barlow	Head	M		76	Pauper	Do Poole	
			Sarah Lockley	Annuit	M		67	Do	Do Winfrith	
46		Do	Hannah Gay	Head	Wid		28	Do	Do Bushy in parish of Corfe Castle	
Total of Houses		46		5	Total of Persons		7	11		

Figure 1: A Page from the 1851 Census Enumerators' Books

```
10/DATA COLLECTION FILE/C1000/K.SCHURER/3 JULY 1984
20/HOUSEHOLD CENSUS/1851
31/6073/STANDARD/RURAL
32/ENGLAND/DORSET/CORFE CASTLE
41/PRO HO 107 1856 130-162
42/ALL THAT PART OF THE PARISH OF CORFE CASTLE WHICH LIES TO THE SOUTH OF +
  CHALLOW HILLS AND EAST OF THE STREAM RUNNING FROM BLASHENWEL TO WEST MILLS +
  INCLUDING WILLIAM SMITHS AND ISAAC PHILLIPS' AT THE BRIDGE REV G.HUBBARD +
  BREAKNECKS CHALLOW SANDY HILLS WOLLGARSTON AILWOOD WESTWOOD TAPERSHILL EAST +
  TOWNSEND AND EAST STREET ''/'' BOTH SIDES BOAR MILL ROUND THE MARKET PLACE +
  AND WEST STREET ''/'' BOTH (SIDES) TO WEST TOWNSEND
.
.
.
81/CORFE CASTLE/DIOCESE OF SALISBURY/((?))/CORFE CASTLE/((?))
50/51/EAST STREET
60/HENRY,DAY/HEAD/MAR/36/-/JOURNEYMAN CARPENTER/DORSET,CORFE CASTLE
60/ELIZABETH,DO/WIFE/MAR/-/31/-/DO,DO
60/GEORGE,DO/SON/-/14/-/SCHOLAR/DO,DO
60/ALBERT,DO/SON/-/8/-/DO/DO,DO
60/MARY SARAH,DO/DAUR/-/-/5/DO/DO,DO
60/ANN,DO/DAUR/-/-/2/-/DO,DO
60/SUSAN,DO/DAUR/-/-/1M/-/DO,DO
70
50/52/ALMS HOUSE EAST STREET
60/MARK,COOK/HEAD/MAR/74/-/PAUPER SHEPHERD/DO,WORTH MATRAVERS
60/MAY,DO/WIFE/MAR/-/73/DO/DO,WINTERBOURNE
70
50/53/DO
60/JANE,SERLEY/HEAD/U/-/56/DO/DO,CORFE CASTLE
89/((LONG LINE APPEARS TO HAVE BEEN DRAWN UNDER ABOVE BY MISTAKE +
  THE ENUMERATOR CORRECTED THIS))
60/WILLIAM,/SON/MAR((?))/34/-/DO/DO,DO
89/((NO DITTO OR SURNAME FOR ABOVE))
70
50/54/DO
60/JOHN,PARKER/HEAD/MAR/44/-/DO (AG LAB)/DO,DO
89/((THE ENUMERATOR ORIGINALLY ENTERED HANDLEY AS THE PARISH OF BIRTH BUT +
  CROSSED IT OUT AND SUBSTITUTED A DITTO))
60/ELIZABETH,DO/WIFE/MAR/-/49/DO/DO,HANDLEY
89/((SHORT LINEDRAWN UNDER THE ABOVE BY MISTAKE LATER CORRECTED))
60/GEORGE,/SON/U/16/-/GENERAL LAB/DO,CORFE CASTLE
89/((NO DITTO OR SURNAME FOR ABOVE))
50/55/DO
60/ELIZABETH,MACLEAN/HEAD/U/-/74/PAUPER/DO,POOLE
60/SARAH,STREKLEY((?))/NURSE/W/-/61/DO/DO,WINFRITH
70
50/56/DO
60/HANNAH,GUY/HEAD/WIR/-/28/DO/DO,BURBY IN THE PARISH OF CORFE CASTLE
70
82/4,,/7,10
89/((THE ENUMERATOR ORIGINALLY ENTERED THE NUMBER OF INHABITED HOUSES AS 6 +
  THEN CHANGED IT TO 5 THEN AGAIN TO 4. ACCORDING TO THE NUMBER OF LONG LINES +
  IT SHOULD BE 5. ALSO THE WHOLE PAGE WAS BRACKETTED IN THE LEFT HAND MARGIN +
  AND ALONG SIDE IT THE WORD BORO ENTERED))
.
.
.
90
```

Figure 2: An Example of the Data Collection File

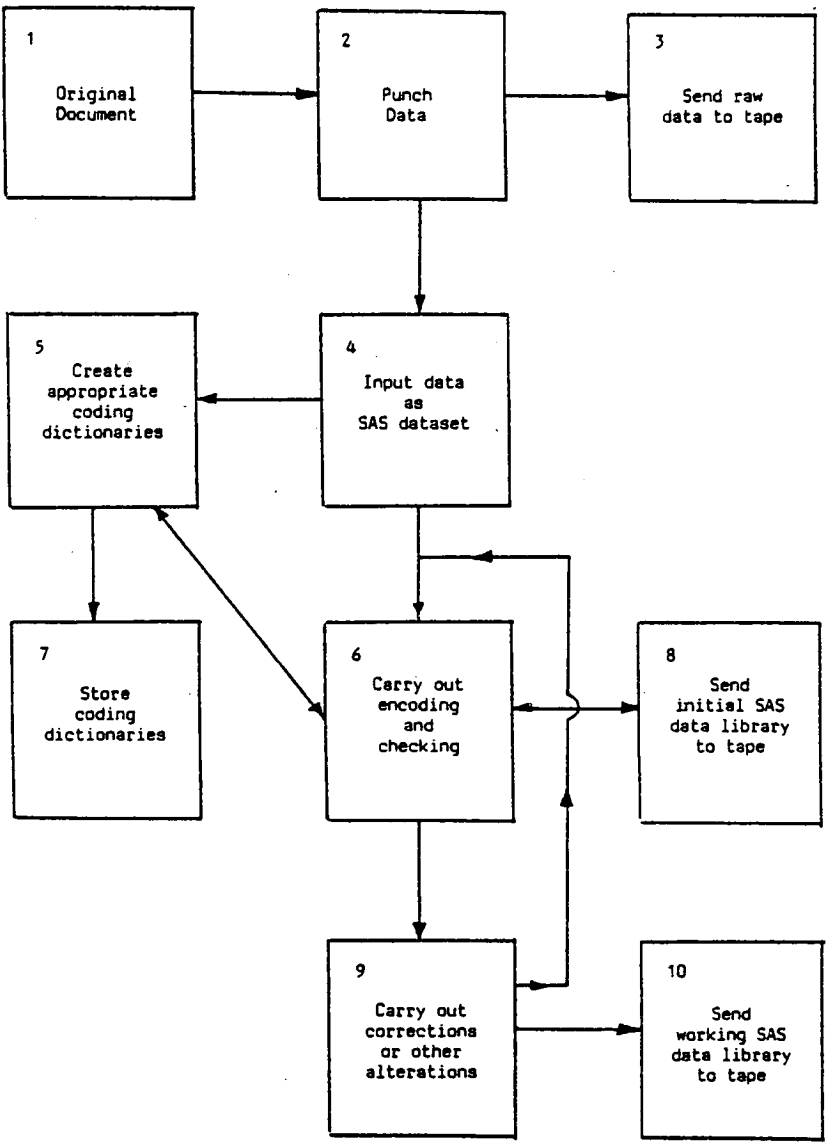


Figure 3: The Input Stage

OBS	PNAME	SNAME	RELAT	COND	OCC	BPCNTY	BPCNTY	BPCNTR	PERSONID	MAGE	FAGE	H	N	HF
1	THOMAS	SCOTT	HEAD	MARR	SHEPARD		NORTHUMBERLAND		1	29	.	1	1	1
2	MARTHA	SCOTT	WIFE	MARR	DRESSMAKER	HAWARDEN	FLINTSHIRE		2	.	24	1	2	1
3	WILLIAM	EDWARDS	HEAD	MAR	BOOT & SHOEMAKER	PENTROBIN	FLINTSHIRE		1	55	.	2	3	2
4	MARY	EDWARDS	WIFE	MARR	BOOT & SHOEMAKERS WIFE	BROUGHTON	FLINTSHIRE		2	.	49	2	4	2
5	GEORGE	EDWARDS	SON	U	BOOT & SHOEMAKER	BROUGHTON	FLINTSHIRE		3	29	.	2	5	2
6	EDWIN	EDWARDS	SON		SCHOLAR	BROUGHTON	FLINTSHIRE		4	13	.	2	6	2
7	EDWD	EDWARDS	SON		SCHOLAR	BROUGHTON	FLINTSHIRE		5	10	.	2	7	2
8	JAMES	EDWARDS	SON		SCHOLAR	BROUGHTON	FLINTSHIRE		6	5	.	2	8	2
9	WILLIAM	WILLIAMS	HEAD	MARR	BRICK MAKER	HAWARDEN	FLINTSHIRE		1	79	.	3	9	3
10	ELIZTH	WILLIAMS	WIFE	MARR		HAWARDEN	FLINTSHIRE		2	.	80	3	10	3
11	JOHN	MACKINTOSH	HEAD	MARR	FARM LABOURER			SCOTLAND	1	32	.	4	11	4
12	MARGARET	MACKINTOSH	WIFE	MARR	FARM LAB WIFE			SCOTLAND	2	.	25	4	12	4
13	ELIZTH	MACKINTOSH	DAUGH			BROUGHTON	FLINTSHIRE		3	.	3	4	13	4
14	WILLIAM	MACKINTOSH	SON			WREXHAM	DENBIGHSHIRE		4	1	.	4	14	4
15	MARRAIT	WILLIAMS	SERVT	U	SERVT	CHESTER	CHESHIRE		5	.	16	4	15	4
16	HANNAH	LANNAH	SERVT	U	SERVT	HOLD	FLINTSHIRE		6	.	29	4	16	4
17	MARY HELLON	HUMPHRYS	VISITOR			LIVERPOOL	LANCASHIRE		7	.	3?	4	17	4
18	DAVID	DAVIES	HEAD	MARR	LABOURER	PENTROROBIN?	FLINTSHIRE		1	63	.	5	18	5
19	ELIZTH	DAVIES	WIFE	MARR	LAB WIFE	DODLESTON	CHESHIRE		2	.	73	5	19	5
20	JOHN	DAVIES	SON	MARR	SHOEMAKER	BROUGHTON	FLINTSHIRE		3	28	.	5	20	5
21	ANN	DAVIES	DAUG IN LAW	MARR	SHOEMAKERS WIFE	WILLASTON	CHESHIRE		4	.	29	5	21	5
22	BERTHA	DAVIES	GRAND DAUGH		SHOEMAKER DAUGH	BROUGHTON	FLINTSHIRE		5	.	1	5	22	5
23	ELIZTH	OWENS	HEAD	WIDOW		HOPE	FLINTSHIRE		1	.	75	6	23	6
24	JOHN	OWEN	SON	UNMARRD	CARTER	HOPE	FLINTSHIRE		2	52	.	6	24	6
25	MARTHA	OWEN	GRAND DAUGH	U	G DAUGH	BROUGHTON	FLINTSHIRE		3	.	15	6	25	6
26	GEORGE	OWEN	GRANDSON		G SON	BROUGHTON	FLINTSHIRE		4	12	.	6	26	6
27	HANNAH	OWENS	GRAND DAU		G DAU	BROUGHTON	FLINTSHIRE		5	.	10	6	27	6
28	JOHN	EDWARDS	SERVT		SERVT	HAWARDEN	FLINTSHIRE		6	38	.	6	28	6
29	RICHD	EDGE	HEAD	MARR	BLACKSMITH	LISNIA?	SHROPSHIRE		1	34	.	7	29	7
30	ELIZTH	EDGE	WIFE	MARR	BLACKSMITHS WIFE	HOLD	FLINTSHIRE		2	.	29	7	30	7
31	MARIA	EDGE	DAUGH		SCHOLAR	HOLD	FLINTSHIRE		3	.	8	7	31	7
32	WILLIAM	EDGE	SON		SCHOLAR	BROUGHTON	FLINTSHIRE		4	6	.	7	32	7
33	SARAH	EDGE	DAUGH			SALTNEY	FLINTSHIRE		5	.	2	7	33	7
34	JOSH	BENION	HEAD	MARR	LABOURER	BURTON GREEN	DENBIGHSHIRE		1	43	.	8	34	8
35	MARGARET	BENION	WIFE	MARR	LAB WIFE	LNYDDYN	DENBIGHSHIRE		2	.	30	8	35	8
36	JOSH	BENION	SON		SCHOLAR	BROUGHTON	FLINTSHIRE		3	9	.	8	36	8
37	MARY	BENION	DAUGH		SCHOLAR	BROUGHTON	FLINTSHIRE		4	.	6	8	37	8
38	EDWD	CANNAH	HEAD	MAR	LABOURER	HOLD	FLINTSHIRE		1	53	.	9	38	9
39	MARIA	CANNAH	WIFE	MAR	LAB WIFE	SORTBOP	FLINTSHIRE		2	.	54	9	39	9
40	EDWD	CANNAH	SON		SCHOLAR	BROUGHTON	FLINTSHIRE		3	12	.	9	40	9

The three variables on the far right H, N, HF are automatically generated by the computer. These form a unique identification number to the household, individual and houseful respectively. The variable Personid numbers individuals within the household, sequentially.

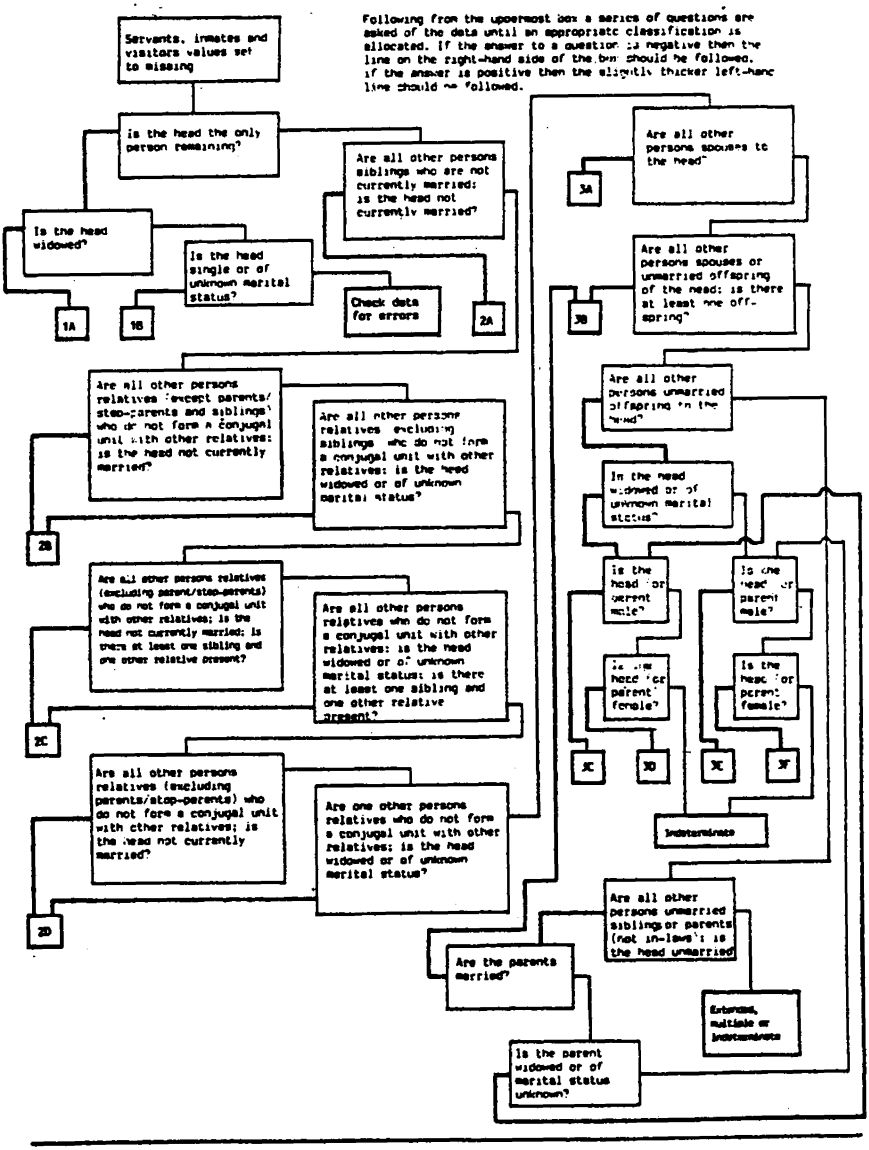
Figure 4: Example of the Data Library Member SPBASE (Broughton, Flintshire – 1851)

TABLE A 1 HOUSEHOLD STRUCTURE: HOUSEHOLDS BY KIN COMPOSITION

Household types	Classes	Number of households	Proportion of all households %
1. Solitaries (singlecons in housenolds)	1a) given as widowed		
	1b) given as non-married or of unknown marital status		
	Sub-total		
2. No family households (co-residents amongst whom no conjugal family unit can be discerned)	2a) co-resident siblings only		
	2b) other co-resident relatives only		
	2c) co-resident siblings and other relative(s)		
	2d) co-residents with no familial relationship given		
	Sub-total		
3. Simple family households (conjugal family units only)	3a) married couples without offspring		
	3b) married couples with offspring		
	3c) widowers with offspring		
	3d) widows with offspring		
	3e) unmarried male parent with illegitimate offspring		
	3f) unmarried female parent with illegitimate offspring		
	Sub-total		
4. Extended family units (conjugal family units having kin-linked individuals)	4a) extension upwards (of which ___ fathers ___ mothers)		
	4b) extension downwards (of which ___ grandchildren only)		
	4c) extension sideways (of which ___ brothers only ___ sisters only)		
	4d) combinations of 4a-4c, or any other form of extension		
	Sub-total		
5. Multiple family households (two or more kin-linked conjugal family units)	5a) households with secondary units disposed upwards (of which ___ also extended)		
	5b) households with secondary units disposed downwards (of which ___ also extended)		
	5c) households with secondary units disposed sideways (of which ___ also extended in other directions)		
	5d) households with secondary units disposed sideways, no member of parental generation present (of which ___ also extended)		
	5e) combinations of 5a-5d, or any other multiple household arrangement (of which ___ also extended)		
	Sub-total		
6. Indeterminate (households where kin linkages are insufficient for classification in any of above)			
	Sub-total		
	TOTAL		100

Definitional and explanatory notes to this table will be found below.
Researcher's comments to be inserted below.

Figure 5: Model Tables for Analysing Families, Households and Housefuls, prepared by the Cambridge Group



'not currently married' = unmarried; widowed; married spouse absent or of unknown marital status
'widowed' = married, spouse absent; separated or divorced
Also note that the system allows for polygamous relationships hence the use of the term 'spouses'

Figure 6: Assignment of Household Classification